

PATENT  
Attorney Docket No.: 16869S-094000US  
Client Ref. No.: W1153-01EJ

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re application of:

HIDEOMI IDEI *et al.*

Application No.: 10/656,096

Filed: September 5, 2003

For: MANAGEMENT SERVER FOR  
ASSIGNING STORAGE AREAS  
TO SERVER, STORAGE  
APPARATUS SYSTEM AND  
PROGRAM

Customer No.: 20350

Examiner: Unassigned

Technology Center/Art Unit: 2141

Confirmation No.: 9922

**PETITION TO MAKE SPECIAL FOR  
NEW APPLICATION UNDER M.P.E.P.  
§ 708.02, VIII & 37 C.F.R. § 1.102(d)**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

This is a petition to make special the above-identified application under MPEP § 708.02, VIII & 37 C.F.R. § 1.102(d). The application has not received any examination by an Examiner.

(a) The Commissioner is authorized to charge the petition fee of \$130 under 37 C.F.R. § 1.17(i) and any other fees associated with this paper to Deposit Account 20-1430.

12/22/2004 CNGUYEN 00000098 201430 10656096  
01 FC:1464 130.00 DA

(b) All the claims are believed to be directed to a single invention. If the Office determines that all the claims presented are not obviously directed to a single invention, then Applicants will make an election without traverse as a prerequisite to the grant of special status.

(c) Pre-examination searches were made of U.S. issued patents, including a classification search and a key word search. The classification search was conducted on or around November 5, 2004 covering Class 709 (subclasses 201, 203, 213, and 214) and Class 711 (subclasses 148 and 170), by a professional search firm, Lacasse & Associates, LLC. The key word search was performed on the USPTO full-text database including published U.S. patent applications. The inventors further provided two references considered most closely related to the subject matter of the present application (see references #8-9 below), which were cited in the Information Disclosure Statements filed on September 5, 2003.

(d) The following references, copies of which are attached herewith, are deemed most closely related to the subject matter encompassed by the claims:

- (1) U.S. Patent No. 6,606,690 B2;
- (2) U.S. Patent Publication No. 2002/0078174 A1;
- (3) U.S. Patent Publication No. 2002/0156984 A1;
- (4) U.S. Patent Publication No. 2003/0110263 A1;
- (5) U.S. Patent Publication No. 2003/0236884 A1;
- (6) U.S. Patent Publication No. 2004/0193827 A1;
- (7) U.S. Patent Publication No. 2004/0194061 A1;
- (8) Japanese Patent Publication No. JP 10-333950; and
- (9) Evaluator Group, Inc., "Virtualization of Disk Storage," The Evaluator Series, WP-0007-1, September 2000, at pp. ii, 1-12.

(e) Set forth below is a detailed discussion of references which points out with particularity how the claimed subject matter is distinguishable over the references.

A. Claimed Embodiments of the Present Invention

The claimed embodiments relate to a system including a management server which manages storage areas of storage apparatuses as virtual storage areas.

Independent claim 1 recites a management server connected to a plurality of servers to manage storage areas included in storage apparatuses as virtual storage areas. The storage apparatuses are shared by the plurality of servers. The storage apparatuses include assignment areas which are storage areas assigned to at least one of the plurality of servers. The management server is responsive to an area assignment instruction of storage areas exceeding unassigned areas received from one of the plurality of servers to release at least part of the assignment areas of other servers as unassigned areas and assign the areas to one of the plurality of servers.

Independent claim 7 recites a storage apparatus system comprising a storage apparatuses; and a management server connected to a plurality of servers and the storage apparatuses. The management server manages storage areas of the storage apparatuses as virtual storage areas. The storage apparatuses are shared by the plurality of servers. The storage apparatuses include assignment areas which are storage areas assigned to at least one of the plurality of servers. The management serve is responsive to an area assignment instruction of storage areas exceeding unassigned areas received from one of the plurality of servers to release at least one of assignment areas of other servers as unassigned area and assign the areas to one of the plurality of servers.

Independent claim 13 recites a computer program product for a management server which manages storage areas included in storage apparatuses as virtual storage areas. The management server is connected to a plurality of servers. The storage apparatuses are shared by the plurality of servers through the management server and include assignment areas which are storage areas assigned to at least one of the plurality of servers. The computer program product comprises code for being responsive to an area assignment instruction of storage areas exceeding unassigned areas received from one of the plurality of servers to release at least part of assignment areas of other servers as unassigned areas and assign the area to one of the plurality of servers; and a computer readable storage medium for storing the code.

One of the benefits that may be derived is that even when the unassigned area is insufficient, the storage area can be assigned to the server issuing the assignment request without waiting until the storage capacity is increased by increase of a new storage apparatus in the SAN or the like. Further, even when the assignment request of the storage areas exceeding the unassigned areas is issued by the server, the storage areas can be assigned to the server to thereby utilize the storage area in the storage pool efficiently.

B. Discussion of the References

None of the following references disclose a management server that is responsive to an area assignment instruction of storage areas exceeding unassigned areas received from one of the plurality of servers to release at least part of the assignment areas of other servers as unassigned areas and assign the areas to one of the plurality of servers.

1. U.S. Patent No. 6,606,690 B2

This reference discloses a system and method for accessing a storage area network as network attached storage. Discussed are SAN servers 302 that communicate with NAS servers 304 using a protocol containing commands that the NAS servers 304 understand. The commands can direct the NAS servers 304 to allocate and deallocate shared storage 110, 112, 114 from the SAN 120 to and from the second network 116. SAN storage manager 404 located in SAN server 302 communicates with NAS server 304 through second network interface 402 to allocate and deallocate storage from NAS server 304. SAN storage manager 404 can send allocation and deallocation directives, and status directives, to NAS server 304. See Figs. 3B and 4; column 5, lines 26-30; column 6, lines 59-62; and column 17, lines 53-57.

2. U.S. Patent Publication No. 2002/0078174 A1

This reference discloses a method and an apparatus for automatically adapting a node in a network. Discussed is a storage management subsystem 2070 that monitors cluster storage usage, supports content rating and pruning, and notifies the content management servers (CMS) 570 of the storage usage of each content provider. Pruning includes freeing the storage used by some blocks or files for use by other blocks or files based on specific usage patterns. The storage management subsystem 2070 includes a

storage management agent that monitors the total local storage availability and the storage availability of a content provider, schedules content pruning, and reports storage usage information and shortage warnings to CMS 570. See Figs. 5 and 20; and paragraphs [0121], [0175], and [0186]. Applicants note that there are 23 known family members for this U.S. Patent application.

3. U.S. Patent Publication No. 2002/0156984 A1

This reference contains the same disclosure as reference #1.

4. U.S. Patent Publication No. 2003/0110263 A1;

This reference discloses a managing storage resources attached to a data network. Discussed is a system for dynamically managing and allocating storage resources 141-143 to application servers 121-123 connected to a data network 100. A managing server 150 identifies all the physical storage resources 141-143 that are connected to network 100 and collects them into a virtual storage pool 160, which is implemented by a plurality of segments that are distributed, using predetermined criteria that are dynamically processed and evaluated, among the physical storage resources, such that the distribution is transparent to each application. Server 150 can re-allocate virtual storage resources to each application according to its actual needs and the level of usage. See Fig. 1; and paragraphs [0041] and [0042].

5. U.S. Patent Publication No. 2003/0236884 A1

This reference discloses a computer system and a method for storage area allocation. Discussed is a device allocation program on a host 11 that sends a request to a storage management server 13. The storage management server 13 compares the storage configurations of the SAN and NAS systems with the presented requirements for the file system area, selects a storage area 12 which best meets the requirements, reconfigures the storage and returns the information about the position of the allocated area to the host 11. See Fig. 1 and paragraph [0100].

6. U.S. Patent Publication No. 2004/0193827 A1

This reference discloses a computer system for managing performances of storage apparatus. In a computer system with a DBMS running thereon, management of the performance of a storage apparatus is executed by using a performance indicator provided by a user job so as to simplify the management of the performance. For this reason, a management server employed in the computer system monitors an operating state of each system element, a response time onto a job and other information. Pre-given information on a process such as a performance requirement the collected monitored information are used by the management server in issuing a command to change allocation of a processing amount to a port, an allocation of a cache area for data, the logical configuration of disc drives and other parameters in order to carry out the new process or in the case where a result of a judgment based on the monitored information indicates that tuning is necessary. In the case of a process for a batch job, a method for estimating a processing time is given to the management server, which issues a setting modification command based on an estimated processing time. In the case of a process for an on-line job, on the other hand, a command to modify settings of a member bearing a heavy load is issued in the case where a response time on a process and/or a throughput do not meet their performance requirement.

7. U.S. Patent Publication No. 2004/0194061 A1;

This reference discloses a method for allocating programs. Disclosed is monitoring that is performed with regard to the performance and capacity of a server executing various business application programs (APs), the performance and capacity of a storage device, or a cluster made up of a plurality of servers. A business service is provided by reallocating the business AP in accordance with these monitoring results. If the load for a particular server has increased, then reallocation can be performed unnecessarily for that business AP. A space scheduler 44 of a private domain management server 11 determines surplus resources. See Fig. 36; and paragraphs [0005], [0006], and [0197].

8. Japanese Patent Publication No. JP 10-333950

This reference discloses a free file space preparation system to secure a free file space and to continue a file preparation processing and a file extension processing by

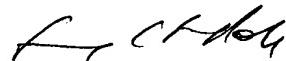
releasing an unused file space in the case of lacking the free file space on an auxiliary storage device at the time of file preparation and file extension. An unused file space release means 105 receives an unused file space release request from a file space securing means 104, finds a file for which the size of logic data 109 is smaller than the size of the file space 108 while retrieving file management information 107 on the auxiliary storage device 106 and releases the unused file space 110 of the difference of the file space 108 and the logic data 109 of the file. When the unused file space 110 is released and the free file space is prepared in the auxiliary storage device 106, control is returned to the file space securing means 104.

9. Evaluator Group, Inc., "Virtualization of Disk Storage," The Evaluator Series, WP-0007-1, September 2000, at pp. ii, 1-12.

This reference discloses the virtualization technique for disk storage. According to the virtualization technique, a management server connected to storage apparatuses and servers using the storage apparatuses manages storage areas of the storage apparatuses connected to the SAN as virtual storage areas (storage pool) collectively and receives requests from the servers to the storage apparatuses. The management server accesses to the storage areas of the storage apparatuses connected thereunder in response to the requests from the servers and returns its results to the servers.

(f) In view of this petition, the Examiner is respectfully requested to issue a first Office Action at an early date.

Respectfully submitted,



Chun-Pok Leung  
Reg. No. 41,405

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, 8<sup>th</sup> Floor  
San Francisco, California 94111-3834  
Tel: 650-326-2400  
Fax: 415-576-0300  
Attachments  
RL:rl  
60362119 v1

W1153

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-333950

(43)Date of publication of application : 18.12.1998

(51)Int.CI. G06F 12/00

(21)Application number : 09-147587

(71)Applicant : NEC SOFTWARE LTD

(22)Date of filing : 05.06.1997

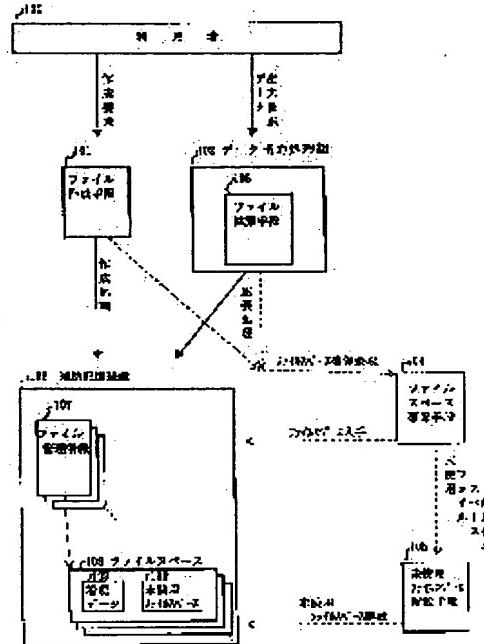
(72)Inventor : YAMAUCHI KOJI

## (54) FREE FILE SPACE PREPARATION SYSTEM

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To secure a free file space and to continue a file preparation processing and a file extension processing by releasing an unused file space in the case of lacking the free file space on an auxiliary storage device at the time of file preparation and file extension.

**SOLUTION:** An unused file space release means 105 receives an unused file space release request from a file space securing means 104, finds a file for which the size of logic data 109 is smaller than the size of the file space 108 while retrieving file management information 107 on the auxiliary storage device 106 and releases the unused file space 110 of the difference of the file space 108 and the logic data 109 of the file. When the unused file space 110 is released and the free file space is prepared in the auxiliary storage device 106, control is returned to the file space securing means 104.



## LEGAL STATUS

[Date of request for examination] 05.06.1997

[Date of sending the examiner's decision of rejection] 03.03.2000

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

## (12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-333950

(43)公開日 平成10年(1998)12月18日

(51)Int.Cl.<sup>6</sup>  
G 0 6 F 12/00識別記号  
5 0 1F I  
G 0 6 F 12/00

5 0 1 M

審査請求 有 請求項の数 5 O.L (全 13 頁)

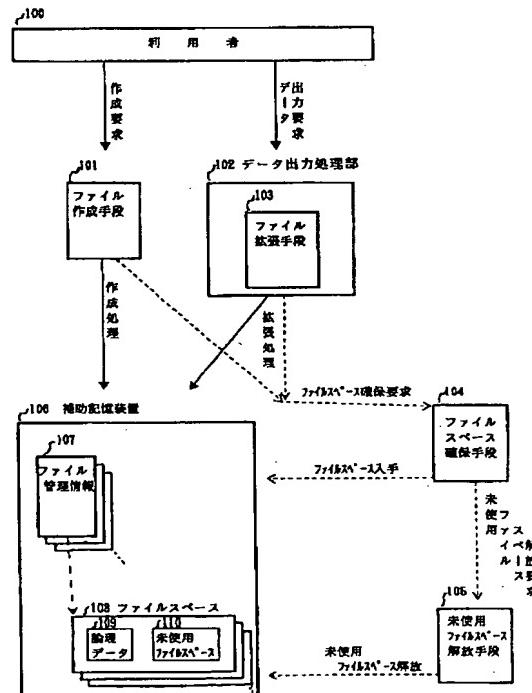
(21)出願番号	特願平9-147587	(71)出願人 000232092 日本電気ソフトウェア株式会社 東京都江東区新木場一丁目18番6号
(22)出願日	平成9年(1997)6月5日	(72)発明者 山内 浩二 東京都江東区新木場一丁目18番6号 日本 電気ソフトウェア株式会社内
		(74)代理人 弁理士 山下 積平

## (54)【発明の名称】 空きファイルスペース作成方式

## (57)【要約】

【課題】 ファイル作成、ファイル拡張時において、補助記憶装置上の空きファイルスペースが不足している場合、ファイルのファイルスペース内の実質使用されていない未使用ファイルスペースを解放することで空きファイルスペースを確保し、補助記憶装置上のファイルスペースを有効利用する空きファイルスペース作成方式を提供する。

【解決手段】 補助記憶装置からスペース管理テーブルを読み込み、スペース管理テーブル内の解放する未使用ファイルスペースに対応した部分を未使用状態にして、スペース管理テーブルを補助記憶装置に書き戻し、磁気ディスク装置からファイルディレクトリを読み込み、ファイルディレクトリ内の位置情報とサイズ情報を未使用ファイルスペースを解放した状態のものに変更して、ファイルディレクトリを補助記憶装置に書き戻す。



## 【特許請求の範囲】

【請求項1】 補助記憶装置上にすでに存在するファイルを調べて、該ファイルのファイルスペースと論理データとの差分である未使用ファイルスペースを見つけ、該未使用ファイルスペースを解放して要求されたサイズの空きファイルスペースを前記補助記憶装置上に作成する未使用ファイルスペース解放手段と、  
要求されたサイズの前記ファイルスペースを前記補助記憶装置上の前記空きファイルスペースから確保し、要求された前記ファイルスペースのサイズが前記空きファイルスペースのサイズよりも大きい場合に前記未使用ファイルスペース解放手段を呼び出すファイルスペース確保手段と、  
利用者からのファイル作成要求を受け取り、作成するファイルのファイル名とファイルサイズとを入力して、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に前記ファイルスペースを確保し、該ファイルスペースにファイルを作成するファイル作成手段と、  
前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に拡張すべき前記ファイルスペースを確保し、ファイル拡張を行うファイル拡張手段と、  
利用者からのデータ出力要求を受け取り、利用者から入力されたデータを前記ファイルスペース内の前記未使用ファイルスペースに書き出し、該未使用ファイルスペースが不足した場合に、前記ファイル拡張手段を呼び出し、前記書き出しを継続するデータ出力手順とを有することを特徴とする空きファイルスペース作成装置。

【請求項2】 前記未使用ファイルスペース解放手段は、前記補助記憶装置からスペース管理テーブルを読み込み、該スペース管理テーブル内の解放する未使用ファイルスペースに対応した部分を未使用状態にして、該スペース管理テーブルを前記補助記憶装置に書き戻し、前記磁気ディスク装置からファイルディレクトリを読み込み、該ファイルディレクトリ内の位置情報とサイズ情報を未使用ファイルスペースを解放した状態のものに変更して、該ファイルディレクトリを前記補助記憶装置に書き戻すことを特徴とする請求項1に記載の空きファイルスペース作成装置。

【請求項3】 補助記憶装置上にすでに存在するファイルを調べて、該ファイルのファイルスペースと論理データとの差分である未使用ファイルスペースを見つけ、該未使用ファイルスペースを解放して要求されたサイズの空きファイルスペースを前記補助記憶装置上に作成する未使用ファイルスペース解放手順と、  
要求されたサイズの前記ファイルスペースを前記補助記憶装置上の前記空きファイルスペースから確保し、要求された前記ファイルスペースのサイズが前記空きファイルスペースのサイズよりも大きい場合に前記未使用ファイルスペース解放手順を呼び出すファイルスペース確保手順と、

利用者からのファイル作成要求を受け取り、作成するファイルのファイル名とファイルサイズとを入力して、前記ファイルスペース確保手順を呼び出して前記補助記憶装置上に前記ファイルスペースを確保し、該ファイルスペースにファイルを作成するファイル作成手順と、  
前記ファイルスペース確保手順を呼び出して前記補助記憶装置上に拡張すべき前記ファイルスペースを確保し、ファイル拡張を行うファイル拡張手順と、  
利用者からのデータ出力要求を受け取り、利用者から入力されたデータを前記ファイルスペース内の前記未使用ファイルスペースに書き出し、該未使用ファイルスペースが不足した場合に、前記ファイル拡張手順を呼び出し、前記書き出しを継続するデータ出力手順とを有することを特徴とする空きファイルスペース作成方式。

【請求項4】 前記未使用ファイルスペース解放手順は、前記補助記憶装置からスペース管理テーブルを読み込み、該スペース管理テーブル内の解放する未使用ファイルスペースに対応した部分を未使用状態にして、該スペース管理テーブルを前記補助記憶装置に書き戻し、前記磁気ディスク装置からファイルディレクトリを読み込み、該ファイルディレクトリ内の位置情報とサイズ情報を未使用ファイルスペースを解放した状態のものに変更して、該ファイルディレクトリを前記補助記憶装置に書き戻すことを特徴とする請求項3に記載の空きファイルスペース作成方式。

【請求項5】 請求項3又は4に記載の空きスペース作成方式をコンピュータに実行させるためのプログラムを記録したことを特徴とする記録媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、補助記憶装置上にファイルを作成したり、被助記憶装置上のファイルを拡張するときの補助記憶装置上のファイルスペース作成方式に関する。

## 【0002】

【従来の技術】 以下、従来のファイルスペース作成方式について説明する。

【0003】 特開平3-138737号公報に記載の大規模ファイル作成装置は、ファイル作成、ファイル拡張時にある補助記憶装置上にファイルスペースが存在しない場合、他の補助記憶装置上に跨ってファイルスペースを確保するものである。

【0004】 また、特開平3-40145号公報に記載の大規模ファイル作成装置は、1つのファイルが複数且つ異機種の二次記憶装置に跨って存在することを可能にするものである。

【0005】 更に、特開平3-24644号公報に記載のマルチボリュームにおけるファイルの自動拡張方式は、複数のボリュームに跨るファイルの動的な自動拡張を可能にするものである。

【0006】更に、特開平1-161548号公報に記載のオンラインファイル拡張方式は、ネットワークに複数のサブシステムが接続されたシステムにおいて、各サブシステムのファイルの格納エリアの拡張を、システム稼働中に行うものである。

【0007】更に、特開昭62-66343号公報に記載の大容量記憶装置上のファイル拡張方式は、大容量記憶装置上のファイルを複数のボリュームに跨って拡張するものである。

#### 【0008】

【発明が解決しようとする課題】従来のいずれのファイルスペース確保方式も、単純にある補助記憶装置上に空きファイルスペースがなくなったら、他の補助記憶装置からファイルスペースを跨って確保するだけで、補助記憶装置上のファイルスペースを有効に利用していない。

【0009】その理由は、補助記憶装置上の見かけの空きファイルスペースが物理的になくなってしまって、実際はその補助記憶装置上に、実ファイルスペースサイズより、論理データサイズが小さく、その差分のファイルスペースが実質的に使用されてない未使用ファイルスペースを持つファイルが多数存在して、その未使用ファイルスペースを合計すると充分に空きファイルスペースが存在する場合においても、従来のファイルスペース確保方式では、単純に他の補助記憶装置上からファイルスペースを確保してしまうということである。

【0010】このようなファイルスペース確保方式では、補助記憶装置上のファイルスペースを有効に利用しないまま、やみくもに他の補助記憶装置上のファイルスペースを物理的に使用していくだけで、それぞれの補助記憶装置上のファイルスペースが有効に利用されない。

【0011】【発明の目的】そこで本発明は、ファイル作成、ファイル拡張時において、補助記憶装置上の空きファイルスペースが不足した場合、ファイルのファイルスペース内の実質使用されていない未使用ファイルスペースを解放することで空きファイルスペースを確保し、ファイル作成処理、ファイル拡張処理を継続させ、補助記憶装置上のファイルスペースを有効利用する空きファイルスペース作成方式を提供することを目的とする。

#### 【0012】

【課題を解決するための手段】本発明による空きファイルスペース作成装置は、補助記憶装置上にすでに存在するファイルを調べて、該ファイルのファイルスペースと論理データとの差分である未使用ファイルスペースを見つけて、該未使用ファイルスペースを解放して要求されたサイズの空きファイルスペースを前記補助記憶装置上に作成する未使用ファイルスペース解放手段と、要求されたサイズの前記ファイルスペースを前記補助記憶装置上の前記空きファイルスペースから確保し、要求された前記ファイルスペースのサイズが前記空きファイルスペースのサイズよりも大きい場合に前記未使用ファイルスペース解放手段を呼び出すファイルスペース確保手段と、利用者からのファイル作成要求を受け取り、作成するファイルのファイル名とファイルサイズとを入力して、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に前記ファイルスペースを確保し、該ファイルスペースにファイルを作成するファイル作成手段と、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に拡張すべき前記ファイルスペースを確保し、ファイル拡張を行うファイル拡張手段と、利用者からのデータ出力要求を受け取り、利用者から入力されたデータを前記ファイルスペース内の前記未使用ファイルスペースに書き出し、該未使用ファイルスペースが不足した場合に、前記ファイル拡張手段を呼出し、前記書き出しを継続するデータ出力処理部とを備えることを特徴とする。

ース解放手段を呼び出すファイルスペース確保手段と、利用者からのファイル作成要求を受け取り、作成するファイルのファイル名とファイルサイズとを入力して、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に前記ファイルスペースを確保し、該ファイルスペースにファイルを作成するファイル作成手段と、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に拡張すべき前記ファイルスペースを確保し、ファイル拡張を行なうファイル拡張手段と、利用者からのデータ出力要求を受け取り、利用者から入力されたデータを前記ファイルスペース内の前記未使用ファイルスペースに書き出し、該未使用ファイルスペースが不足した場合に、前記ファイル拡張手段を呼出し、前記書き出しを継続するデータ出力処理部とを備えることを特徴とする。

【0013】また、本発明による空きファイルスペース作成装置は、前記未使用ファイルスペース解放手順は、前記補助記憶装置からスペース管理テーブルを読み込み、該スペース管理テーブル内の解放する未使用ファイルスペースに対応した部分を未使用状態にして、該スペース管理テーブルを前記補助記憶装置に書き戻し、前記磁気ディスク装置からファイルディレクトリを読み込み、該ファイルディレクトリ内の位置情報とサイズ情報を未使用ファイルスペースを解放した状態のものに変更して、該ファイルディレクトリを前記補助記憶装置に書き戻すことを特徴とする。

【0014】本発明による空きファイルスペース作成方式は、補助記憶装置上にすでに存在するファイルを調べて、該ファイルのファイルスペースと論理データとの差分である未使用ファイルスペースを見つけ、該未使用ファイルスペースを解放して要求されたサイズの空きファイルスペースを前記補助記憶装置上に作成する未使用ファイルスペース解放手順と、要求されたサイズの前記ファイルスペースを前記補助記憶装置上の前記空きファイルスペースから確保し、要求された前記ファイルスペースのサイズが前記空きファイルスペースのサイズよりも大きい場合に前記未使用ファイルスペース解放手順を呼び出すファイルスペース確保手順と、利用者からのファイル作成要求を受け取り、作成するファイルのファイル名とファイルサイズとを入力して、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に前記ファイルスペースを確保し、該ファイルスペースにファイルを作成するファイル作成手順と、前記ファイルスペース確保手段を呼び出して前記補助記憶装置上に拡張すべき前記ファイルスペースを確保し、ファイル拡張を行なうファイル拡張手順と、利用者からのデータ出力要求を受け取り、利用者から入力されたデータを前記ファイルスペース内の前記未使用ファイルスペースに書き出し、該未使用ファイルスペースが不足した場合に、前記ファイル拡張手順を呼出し、前記書き出しを継続するデータ出力手順とを有することを特徴とする。

【0015】また、本発明による空きファイルスペース作成方式は、前記未使用ファイルスペース解放手順は、前記補助記憶装置からスペース管理テーブルを読み込み、該スペース管理テーブル内の解放した未使用ファイルスペースに対応した部分を未使用状態にして、該スペース管理テーブルを前記補助記憶装置に書き戻し、前記磁気ディスク装置からファイルディレクトリを読み込み、該ファイルディレクトリ内の位置情報とサイズ情報を未使用ファイルスペースを解放した状態のものに変更して、該ファイルディレクトリを前記補助記憶装置に書き戻すことを特徴とする。

【0016】本発明による記録媒体は、上記の空きスペース作成方式をコンピュータに実行させるためのプログラムを記録したことを特徴とする。

【0017】【作用】本発明においては、ファイル作成及びファイル拡張時に補助記憶装置上に空きファイルスペースがない場合、補助記憶装置上にすでに存在するファイルを調べて、そのファイルのファイルスペースサイズより実際の論理データサイズが小さい場合に存在するそれらの差分である未使用ファイルスペースを見つけ、その未使用ファイルスペースを解放して確保すべきファイルスペースの空きを補助記憶装置上に作成する。これにより必要サイズのファイルスペースが作成できれば、従来、ファイル作成及びファイル拡張で空きファイルスペースがなくて異常終了していたり、実質的に空きが補助記憶装置上に存在するのに、やみくもに他補助記憶装置上へファイルスペースを確保していたケースが救済される。

#### 【0018】

【発明の実施の形態】次に、本発明の第一の実施の形態について図面を参照して詳細に説明する。

【0019】図1を参照すると、本実施形態はファイル作成手段101とデータの出力動作を行うデータ出力処理部102とデータ出力処理部102内のファイル拡張手段103とファイルスペース確保手段104と未使用ファイルスペース解放手段105と補助記憶装置106と補助記憶装置106内に確保されるファイルスペース108とファイルスペース108を管理するファイル管理情報107とファイルスペース108内の論理データ109及び未使用ファイルスペース110とから構成される。

【0020】次に、本発明の第一の実施の形態の動作について、図1を参照して詳細に説明する。

【0021】ファイル作成手段101は、利用者100からファイル作成要求を受け取ったら、指定されたファイルサイズ分のファイルスペース108を補助記憶装置106上に確保するため、ファイルスペース確保手段104へ制御を渡す。制御が戻りファイルスペース108が確保された場合、利用者100から入手したファイル名と確保したファイルスペース108を元にファイル管

理情報107のイメージを作成して、補助記憶装置106上へ書き出す。

【0022】データ出力処理部102は、利用者100からの既存ファイルへのデータ出力要求を受け取り、データ出力するファイルのファイル管理情報107を補助記憶装置106から読み込む。読み込んだファイル管理情報107を元に出力要求されるデータ量が現在のファイルスペース108の領域以内に収まれば、補助記憶装置106に対して実際のデータ出力動作を行う。出力要求されるデータ量が現在のファイルスペース108に収まりきらない場合は、足りないサイズのファイルスペース108を確保するため、ファイル拡張手段103へ制御を渡す。制御が戻りファイル拡張ができた場合は、補助記憶装置106に対して実際のデータ出力動作を行う。

【0023】ファイル拡張手段103は、データ出力処理部102からファイル拡張要求を受け取り、要求されたサイズのファイルスペース108を補助記憶装置106に確保するため、ファイルスペース確保手段104へ制御を渡す。制御が戻りファイルスペース108が確保できた場合は、確保したファイルスペース108の情報を補助記憶装置106上の既存のファイル管理情報107へ附加して書き出して反映し、データ出力処理部102へ制御を戻す。

【0024】ファイルスペース確保手段104は、ファイル作成手段101、ファイル拡張手段103からのファイルスペース確保要求を受け取り、要求ファイルサイズ分の空きファイルスペース108が補助記憶装置106上に存在するかをチェックする。要求ファイルサイズ分の空きが存在するならば、そのサイズ分のファイルスペース108を補助記憶装置106から確保し、ファイル作成手段101、ファイル拡張手段103へ制御を戻す。要求ファイルサイズ分の空きが存在しないならば、補助記憶装置106上に存在するファイルの未使用ファイルスペース110を解放するため、未使用ファイルスペース解放手段105へ制御を渡す。制御が戻り未使用ファイルスペース110が解放されて要求ファイルサイズ分の空きが確保できた場合は、そのサイズ分のファイルスペース108を補助記憶装置106から確保し、ファイル作成手段101、ファイル拡張手段103へ制御を戻す。

【0025】未使用ファイルスペース解放手段105は、ファイルスペース確保手段104からの未使用ファイルスペース解放要求を受け取り、補助記憶装置106上のファイル管理情報107を検索していきながら、ファイルスペース108のサイズより論理データ109のサイズが小さいファイルを見つけ、そのファイルのファイルスペース108と論理データ109の差分の未使用ファイルスペース110を解放していく。未使用ファイルスペース110を解放していく、空きファイルスペー

スを補助記憶装置106に作成したら、ファイルスペース確保手段104へ制御を戻す。

【0026】

【実施例】次に、本発明の実施例について、図面を参照して詳細に説明する。

【0027】図2を参照すると、本実施例はファイル作成手段101とデータの出力動作を行うデータ出力処理部102とデータ出力処理部102内のファイル拡張手段103とファイルスペース確保手段104と未使用ファイルスペース解放手段105と磁気ディスク装置111と磁気ディスク装置111上の使用スペース情報を管理するスペース管理テーブル113と磁気ディスク装置111内に確保されるファイルスペース108とファイルスペース108を管理するファイルディレクトリ112とファイルスペース108内の論理データ109及び未使用ファイルスペース110とから構成される。

【0028】図3は、ファイル作成手段101の利用者100からファイル作成要求を受け取った時の処理の流れを示す図である。

【0029】図4は、利用者100から既存ファイルへのデータ出力要求を受け取った時のデータ出力処理部102の処理の流れを示す図である。

【0030】図5は、データ出力処理部102においてファイル拡張が必要となった時のファイル拡張手段103の処理の流れを示す図である。

【0031】図6は、ファイル作成手段101又はファイル拡張手段103からファイルスペース確保要求を受け取った時のファイルスペース確保手段104の処理の流れを示す図である。

【0032】図7は、ファイルスペース確保手段104から未使用ファイルスペース110の解放要求を受け取った時の未使用ファイルスペース解放手段105の処理の流れを示す図である。

【0033】図3を参照してファイル作成手段101の処理手順を説明する。利用者100からのファイル作成要求を受け取り、指定されたファイル名とファイルサイズ等を入手する(301)。入手したファイルサイズ分のファイルスペース108を磁気ディスク装置111から確保するため、ファイルスペース確保手段104へ制御を渡す(302)。ファイルスペース確保手段104から制御が戻ってきたら、ファイルスペース108が確保できたかをファイルスペース確保手段104のリターン情報からチェックする(303)。ファイルスペース108が確保出来なかった場合は、利用者100にファイル作成が失敗したという異常通知を行う(304)。ファイルスペース108が確保できた場合は、ファイルディレクトリ112のイメージを作成して、利用者100から入手したファイル名と確保したファイルスペース108の位置及びサイズ等をファイルディレクトリ112上に格納する(305)。次にそのファイルディレク

トリ112を磁気ディスク装置111上に書き出す(306)。

【0034】図4を参照してデータ出力処理部102の処理手順を説明する。利用者100から既存ファイルへのデータ出力要求を受け取り、データ出力するファイルのファイルディレクトリ112を磁気ディスク装置111から読み込む(401)。ファイルディレクトリ112内のファイルスペース108の位置とサイズから現在のファイルスペース108の領域外へデータを出力する要求かをチェックする(402)。現在のファイルスペース108の領域内へのデータ出力要求であれば、磁気ディスク装置111に対して実際のデータ出力動作を行う(406)。現在のファイルスペース112の領域外へのデータ出力要求であれば、足りない領域のファイルサイズ分のファイルスペース108を確保するため、ファイル拡張手段103へ制御を渡す(403)。ファイル拡張手段103から制御が戻ってきたら、ファイル拡張が正常に行われたかをリターン情報からチェックする(404)。ファイル拡張に失敗した場合は、利用者100にデータ出力処理部102が失敗したという異常通知を行う(405)。ファイル拡張が成功した場合は、磁気ディスク装置111に対して実際のデータ出力動作を行う(406)。

【0035】図5を参照してファイル拡張手段103の処理手順を説明する。データ出力処理部102からファイル拡張要求を受け取り、足りないサイズのファイルスペース108を確保するため、ファイルスペース確保手段104へ制御を渡す(501)。ファイルスペース確保手段104から制御が戻ってきたら、ファイルスペース108が確保できたかをファイルスペース確保手段104のリターン情報からチェックする(502)。ファイルスペース108が確保できなかった場合は、データ出力処理部102へファイル拡張処理が失敗したことを探る情報をとして通知する(503)。ファイルスペース108が確保できた場合は、新たに確保したファイルスペース108の位置とサイズをファイルディレクトリ112内の既に格納されている現在のファイルスペース108の位置とサイズに付加する形で格納し(504)、磁気ディスク装置111へファイルディレクトリ112を書き出す(505)。次にデータ出力処理部102へファイル拡張処理が成功したことを探る情報をとして通知する(506)。

【0036】図6を参照してファイルスペース確保手段104の処理手順を説明する。ファイル作成手段101又はファイル拡張手段103からのファイルスペース確保要求を受け取り、磁気ディスク装置111からスペース管理テーブル113を読み込む(601)。要求ファイルサイズ分の空きファイルスペース108が磁気ディスク装置111上に存在するかを読み込んだスペース管理テーブル113からチェックし(602)、要求ファ

イルサイズ分の空きが存在するならば、スペース管理テーブル113上の要求ファイルサイズ分のファイルスペース108の領域を使用中状態に変更して、ファイルスペース108を確保し(606)、スペース管理テーブル113を磁気ディスク装置111へ書き出す(607)。次にファイルスペース確保手段104を呼び出した側へリターン情報としてファイルスペース108の確保が成功したこと通知する(608)。要求ファイルサイズ分の空きが存在しないならば、磁気ディスク装置111上に存在するファイルスペース108内の未使用ファイルスペース110を解放するため、未使用ファイルスペース解放手段105へ制御を渡す(603)。未使用ファイルスペース解放手段105から制御が戻ってきたら、要求ファイルサイズ分の空きファイルスペース108が確保できたかを未使用ファイルスペース解放手段105のリターン情報からチェックする(604)。要求ファイルサイズ分の空きファイルスペース108が確保出来なかった場合は、リターン情報としてファイルスペース108の確保が失敗したことをファイルスペース確保手段104を呼び出した側(ファイル作成手段101又はファイル拡張手段103)へ通知する(605)。要求ファイルサイズ分の空きファイルスペース108が確保された場合は、スペース管理テーブル113上の要求ファイルサイズ分のファイルスペース108の領域を使用中状態に変更して、ファイルスペース108を確保し(606)、スペース管理テーブル113を磁気ディスク装置111へ書き出す(607)。次にファイルスペース確保手段104を呼び出した側(ファイル作成手段101又はファイル拡張手段103)へリターン情報としてファイルスペース108の確保が成功したことを通知する(608)。

【0037】図7を参照して未使用ファイルスペース解放手段105の処理手順を説明する。ファイルスペース確保手段104からの未使用ファイルスペース解放要求を受け取り、磁気ディスク装置111から最初に存在するファイルのファイルディレクトリ112を読み込む(701)。次に、読み込んだファイルディレクトリ112内のファイル使用中表示が立っているかをチェックする(702)。ファイル使用中表示が立っている場合は、磁気ディスク装置111上の全てのファイルを調査したかの処理へ制御を移す(710)。ファイル使用中表示が立っていない場合は、ファイルディレクトリ112を元にファイルスペース108内のデータを磁気ディスク装置111から読み込んで論理データ109のサイズを調べ(703)、ファイルスペース108のサイズより論理データ109のサイズが小さく、その差分が未使用ファイルスペース110となっているかをチェックする(704)。未使用ファイルスペース110となっていない場合は、磁気ディスク装置111上の全てのファイルを調査したかの処理へ制御を移す(710)。未

使用ファイルスペース110となっている場合は、磁気ディスク装置111からスペース管理テーブル113を読み込み(705)、読み込んだスペース管理テーブル113に対して、未使用ファイルスペース110のサイズ分のファイルスペース領域を未使用状態へ変更し(706)、先に読み込んでいるファイルディレクトリ112内のファイルスペース108の位置とサイズに解放した未使用ファイルスペース110の情報を反映する(707)。そして、磁気ディスク装置111へファイルディレクトリ112とスペース管理テーブル113を書き出して未使用ファイルスペース110を解放し(708)、そして要求ファイルサイズ分の空きファイルスペース108が磁気ディスク装置111上に作成されたかをチェックする(709)。確保すべき空きが作成された場合は、ファイルスペース確保手段104へリターン情報として要求ファイルサイズ分の空きファイルスペース108の作成が成功したことを通知する(712)。確保すべき空きが作成されなかった場合は、磁気ディスク装置111上の全てのファイルを調査したかをチェックする(710)。全てのファイルを調査した場合は、ファイルスペース確保手段104へリターン情報として要求ファイルサイズ分の空きファイルスペース108の作成が失敗したことを通知する(713)。まだ、全てのファイルを調査していない場合、磁気ディスク装置111から次に存在するファイルのファイルディレクトリ112を読み込み(711)、読み込んだファイルディレクトリ112内のファイル使用中表示が立っているかのチェック処理へ制御を移す(702)。702以降は前述の動作を繰り返す。

#### 【0038】

【発明の効果】以上説明したように本発明によれば、ファイル作成はファイル拡張時に、補助記憶装置上に要求ファイルスペース分の空きが存在しない場合、未使用ファイルスペース解放手段が、実ファイルサイズと実際の論理データサイズとの差分である未使用ファイルスペースを空きファイルスペースとして解放するので、未使用ファイルスペース解放手段により新たに必要になったサイズの空きファイルスペースが作成できれば、補助記憶装置上に要求ファイルスペース分の空きが存在しなくて異常終了していたケースが救済され、補助記憶装置上のファイルスペースが有効に利用できる。

#### 【図面の簡単な説明】

【図1】本発明の第一の実施形態を示す構成図である。

【図2】本発明の実施例を示す構成図である。

【図3】本発明によるファイル作成手段の処理を示す流れ図である。

【図4】本発明によるデータ出力処理部の処理を示す流れ図である。

【図5】本発明によるファイル拡張手段の処理を示す流れ図である。

【図6】本発明によるファイルスペース確保手段の処理を示す流れ図である。

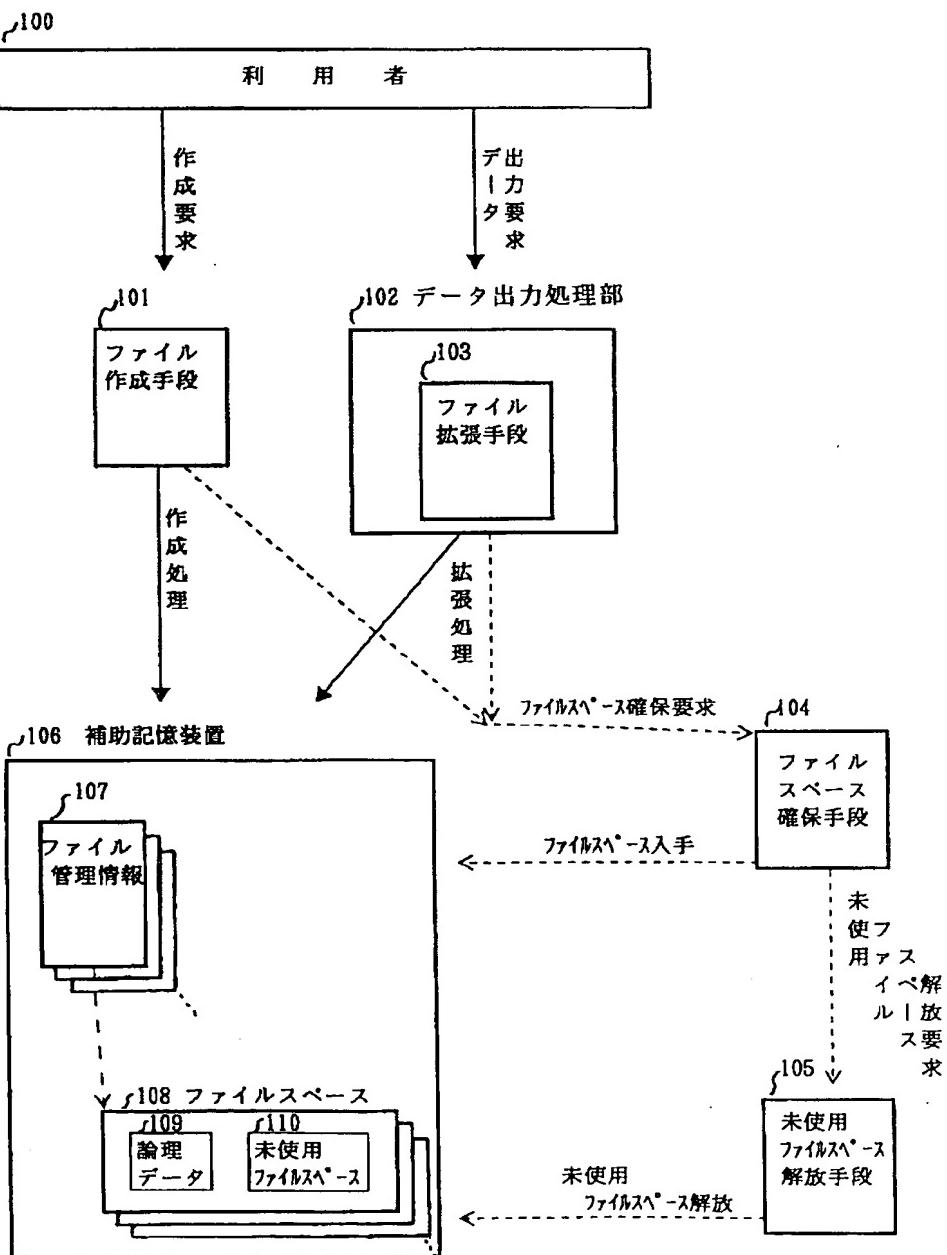
【図7】本発明による未使用ファイルスペース解放手段の処理を示す流れ図である。

【符号の説明】

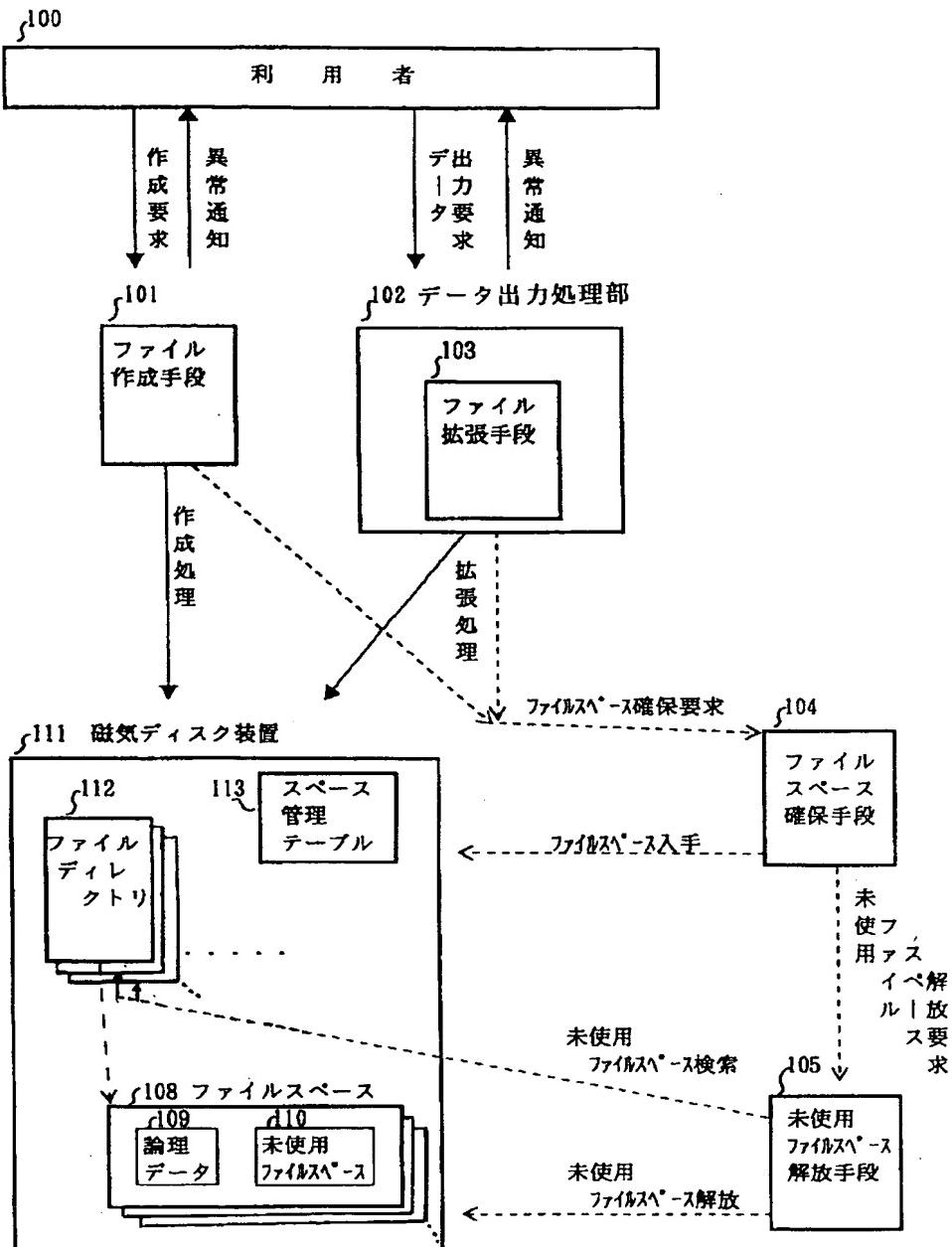
- |     |              |
|-----|--------------|
| 100 | 利用者          |
| 101 | ファイル作成手段     |
| 102 | データ出力処理部     |
| 103 | ファイル拡張手段     |
| 104 | ファイルスペース確保手段 |

- |     |                 |
|-----|-----------------|
| 105 | 未使用ファイルスペース解放手段 |
| 106 | 補助記憶装置          |
| 107 | ファイル管理情報        |
| 108 | ファイルスペース        |
| 109 | 論理データ           |
| 110 | 未使用ファイルスペース     |
| 111 | 磁気ディスク装置        |
| 112 | ファイルディレクトリ      |
| 113 | スペース管理テーブル      |

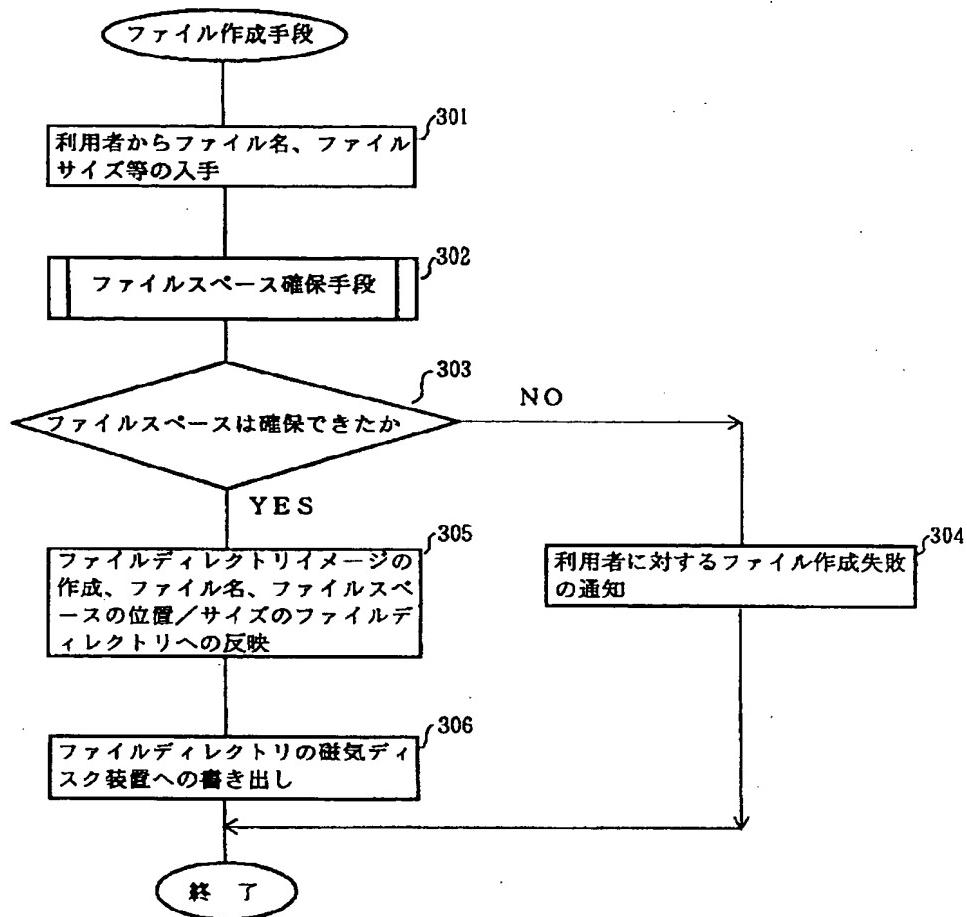
【図1】



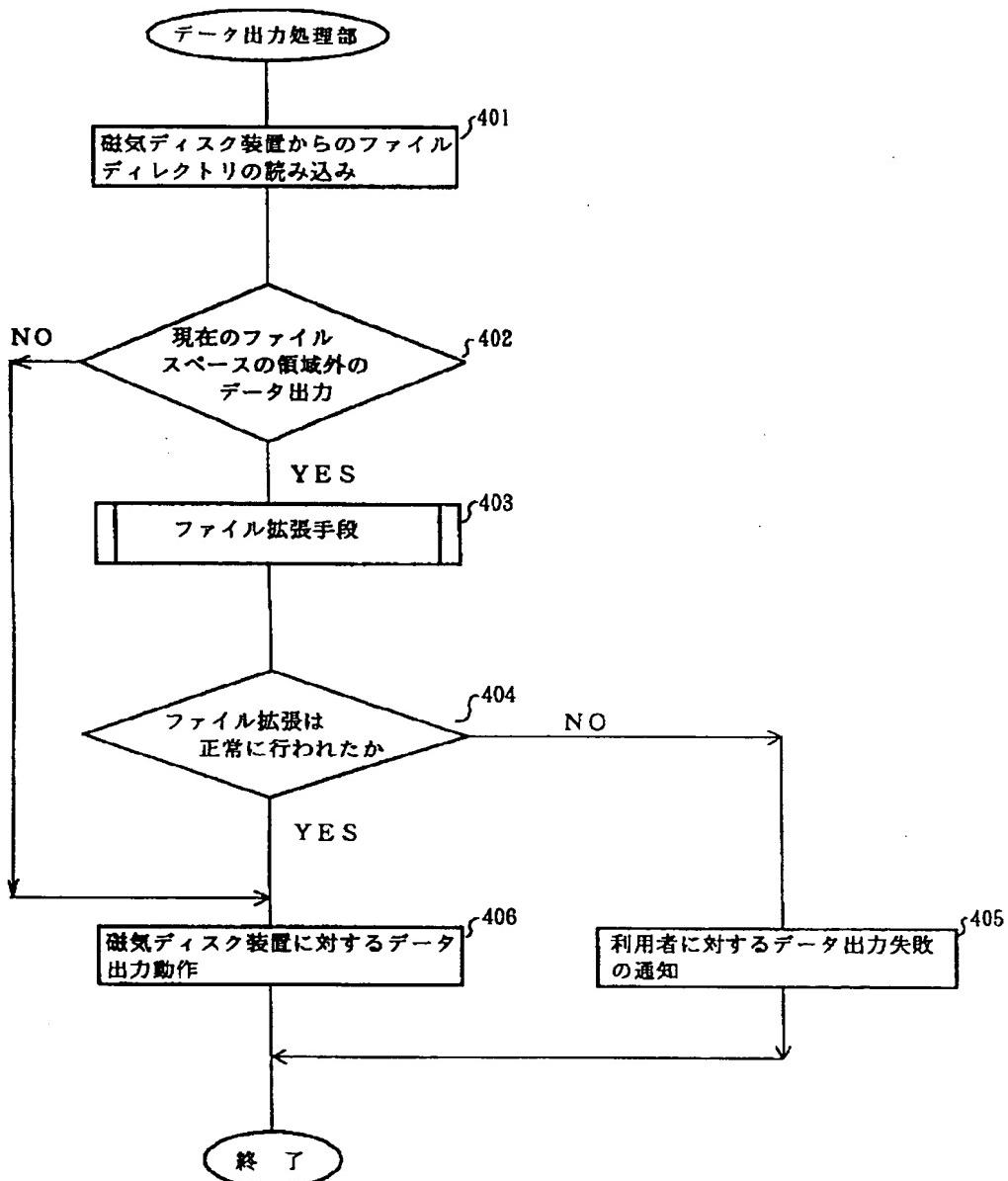
【図2】



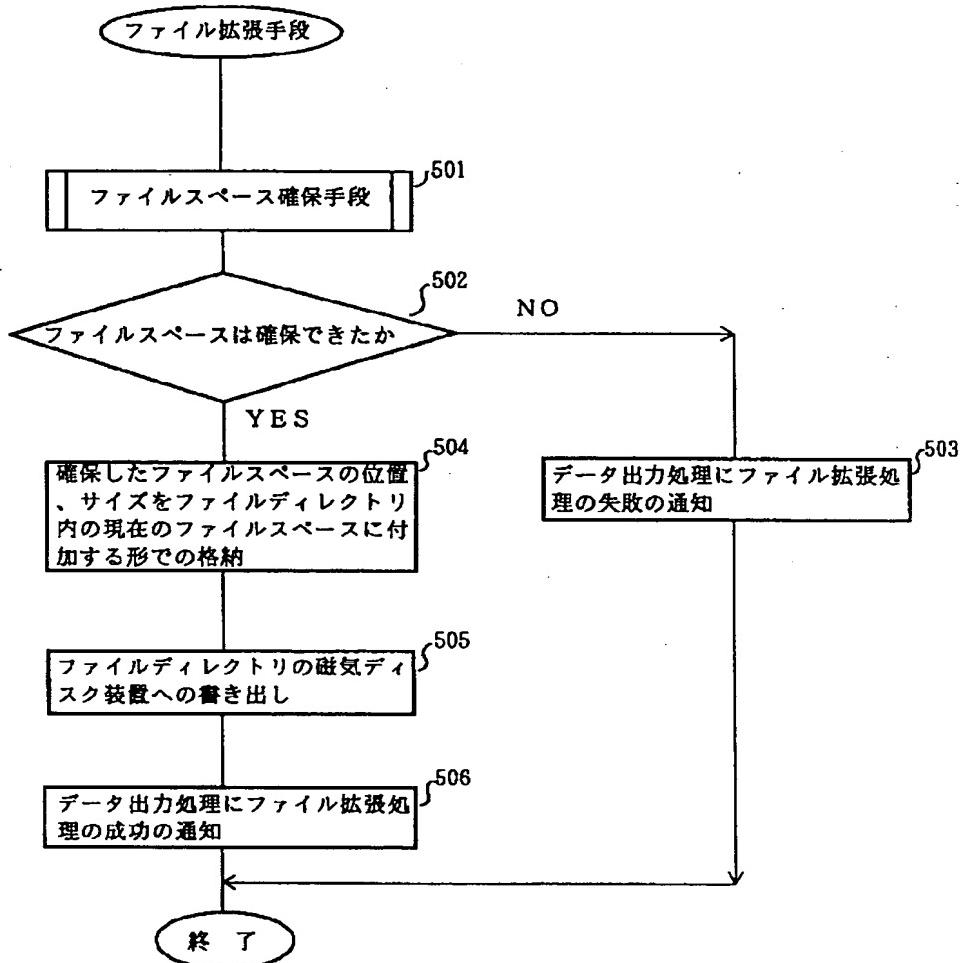
【図3】



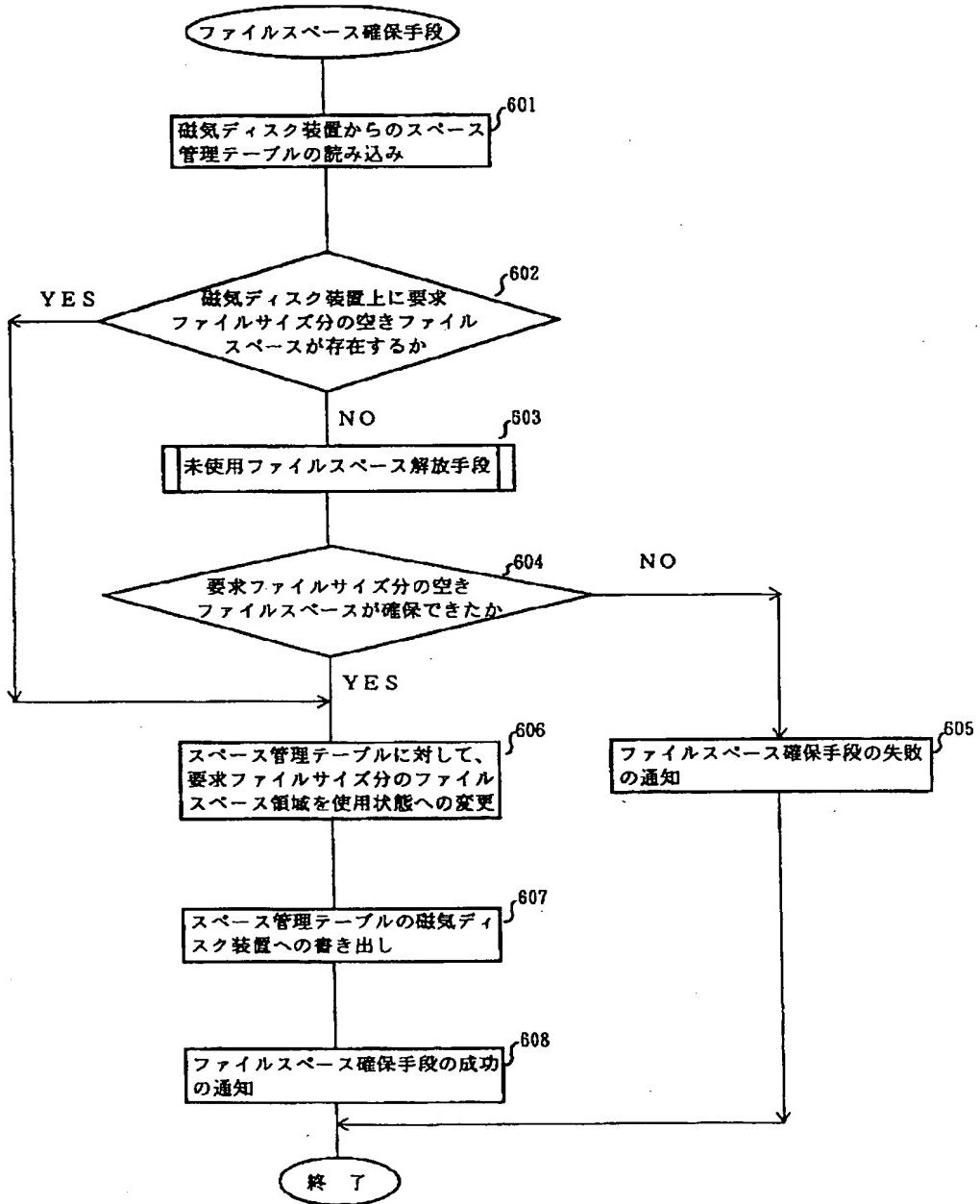
【図4】



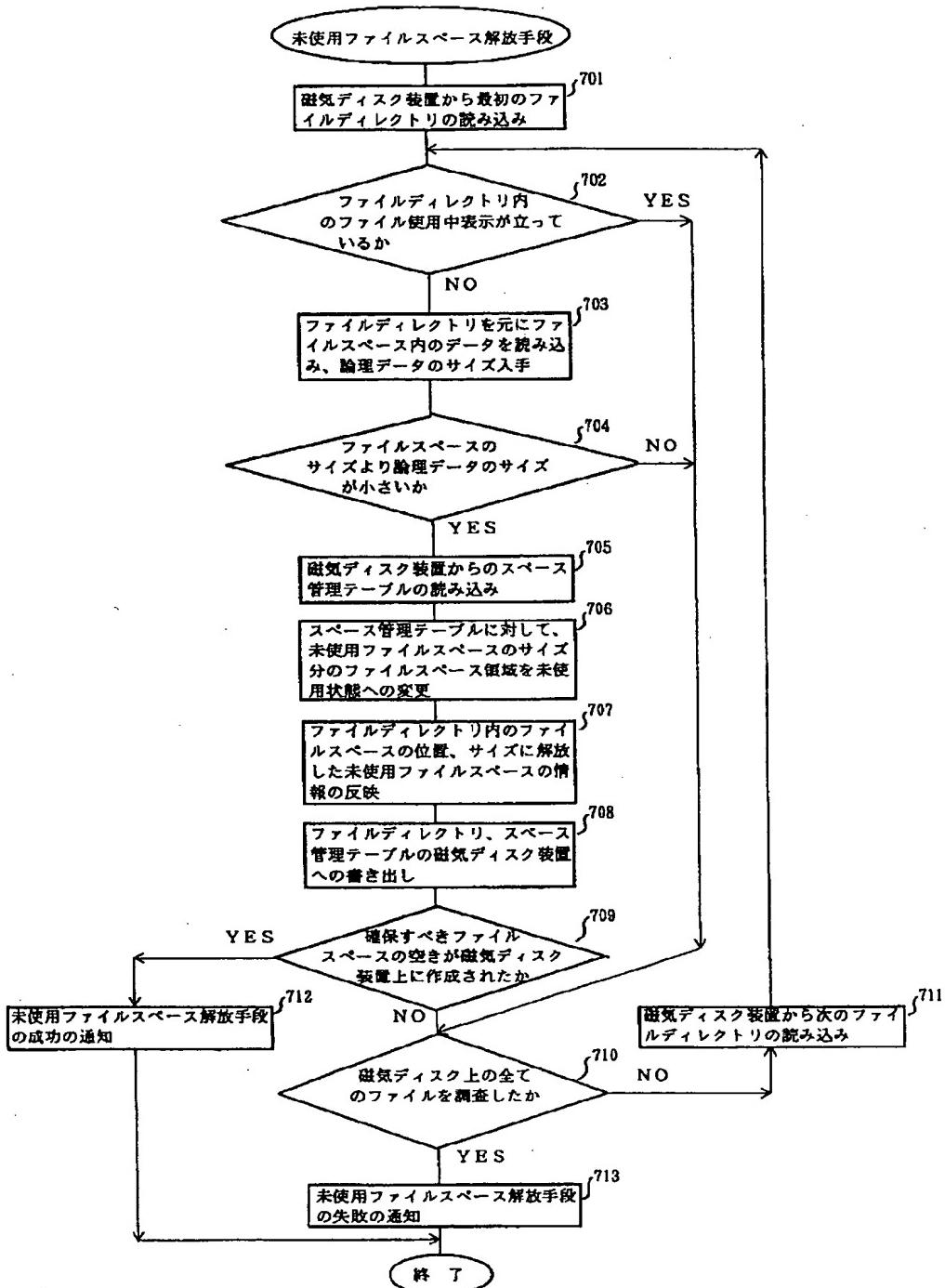
【図5】



【図6】



【図7】



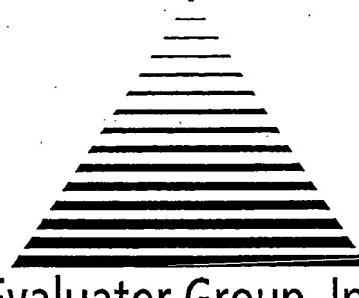
W1153

# ***The Evaluator Series™***

## ***Virtualization of Disk Storage***

**WP-0007-1**

**September 2000**



**Evaluator Group, Inc.**

TM

Comprehensive Information Storage Evaluation

7720 East Belleview Avenue • Suite 210 • Englewood CO 80111

Tel: (303) 221-7867 • Fax: (303) 221-1615

[www.evaluatorgroup.com](http://www.evaluatorgroup.com)

## **Virtualization of Disk Storage (WP-0007-1)**

Copyright 2000 Evaluator Group, Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or stored in a database or retrieval system for any purpose without the expressed written permission of the Evaluator Group.

The information contained in this document is subject to change without notice. The Evaluator Group assumes no responsibility for any errors that may appear.

The Evaluator Group makes no expressed or implied warranties in this document relating to the use or operation of the products described herein.

In no event shall the Evaluator Group be liable for any indirect, special, consequential, or incidental damages arising out of or associated with any aspect of this publication, even if advised of the possibility of such damages.

*The Evaluator Series* is a trademark of the Evaluator Group.

All other trademarks are the property of their respective companies.

## **Virtualization of Disk Storage (WP-0007-1)**

### **Introduction**

Virtualization of disk storage has been in the spotlight recently due primarily to the development of Storage Area Networks. Vendor announcements and papers, as well as press coverage, have described virtualization and characteristics in various ways. The importance and emphasis is well founded given the benefits being promised with the various solutions. However virtualization of disk storage is not new. There have been prior solutions that have provided characteristics that would have to be described as storage virtualization. The most successful of these were the two virtualized disk storage subsystems developed by StorageTek and sold by both StorageTek and IBM and Hewlett-Packard with the AutoRaid storage system. Virtualization of disk storage has been an evolutionary process in what vendors have been providing. An enterprise-wide storage virtualization will be revolutionary by changing the fundamental rules of how we manage storage.

There has been great latitude taken in the description of what virtualization of disk storage is and the characteristics it embodies. This paper will explain what virtualization of disk storage is from a general sense without the specific product differentiated features accentuated by some vendors.

The term virtualization when it pertains to disk storage is an abstraction of storage where the representation of a storage unit to the operating system and applications on a server is divorced from the actual physical storage where the information may be contained. This implies there is some method for providing the translation between the logical and physical. Where and how this is done is the big difference between solutions being provided (or proposed) by vendors. The abstraction, or virtualization, enables many new benefits and capabilities that have been highlighted as part of the benefits with Storage Area Networks. Enabling the ability to deal with storage in abstraction where the details of configuration, device specific features and characteristics are transparent to the server software allows use of differing types of storage, commonality of management tools, and exploitation of a common set of features.

An analogy for the virtualization of disk storage that is illustrative for those that remember or have studied the evolution of computer systems is when the transition was made from real memory to virtual memory in processors. Prior to the implementation of virtual address translation, application programs had to manage their memory space. If a program was going to use more memory than available, mechanisms such as overlays had to be used. With virtual memory, the issues for programming changed dramatically and enabled new capabilities for execution of programs. The change to virtualization of disk storage has the potential to be as dramatic in the management of disk storage.

## **The Benefits of Implementing Virtualization**

Virtualizing disk storage has many obvious benefits that have been extolled and probably a significant number that have yet to be exploited. Understanding the issues for disk storage and how virtualization provides improvements is one way to put perspective and understanding into these benefits.

### **Challenges for IT today with disk storage**

The growth of data storage has to be the most obvious and pressing problem facing IT organizations. The growth being experienced is unprecedented and has led to a multitude of specific issues. Obvious problems are the costs associated with the physical purchase of more storage, the physical space and environmental costs, and the operational time to bring storage into service. The longer-term challenge is the added administration cost for that storage. With the continued growth, this becomes a compounding problem.

This growth of storage exacerbates another severe problem: the availability of people with the skills to manage storage. With a given set of tools, a storage administrator can manage a specific amount of storage. This means that with storage growth, more administrators will be required. This cost is quantifiable but may be moot if there are not enough skilled people to get the job done.

Significant changes in technology as represented in products and infrastructure have occurred over a short period of time and the speed with which these changes are occurring is accelerating. IT organizations struggle to understand and evaluate these new technologies and products. Usually, they do not have the staff to evaluate and form a strategy for deployment of new technology and increasingly are looking to independent organizations to provide information.

The power of information is becoming more visible to executives who see IT organizations as a corporate competitive advantage. This brings a spotlight to IT and a demand on implementing new technologies to remain or increase competitiveness. Making the "correct" strategic decision is critical to companies and to careers.

The increased competitiveness and the amount of information have led to very complex operational environments. There is more demand for "data integration" where information from multiple sources and systems, in differing formats, can be correlated and presented in a unified, informative manner. This expectation is today and is immediate. Achieving it requires significant planning and work.

## **Virtualization of Disk Storage (WP-0007-1)**

All these challenges result in a bottom line issue of cost. They represent significant costs in the administration of storage systems and in the infrastructure implementation to be able to meet increased demands.

### **Virtualization benefits for these challenges**

Accommodating the growth of storage requires that storage can be non-disruptively added and data can be migrated as necessary without operational impact. This means that the storage infrastructure has to be capable of allowing that and the operating system and applications on servers can transparently handle the addition of storage. Storage Area Networks promise the infrastructure for this but it is the virtualization or abstraction of storage that allows it to be transparent and non-disruptive. By dealing with devices from a virtual standpoint, the systems and applications can be shielded from the physical changes that occur to add the storage. The movement of data to balance workloads can be done oblivious to normal operations. Virtualization allows policy-based management of storage to be implemented in a practical manner.

Virtualization allows the management tools to be consistent across varying devices. This means that storage administrators would have to learn only one set of tools to manage storage and that purchasing decisions would be based on costs, reliability, etc. at the planned time of need rather than a decision for a predetermined vendor set done at an earlier time and driven by operational impact required to change. By utilizing a common set of tools, the training would be more focused and require less time and less opportunity for mistakes over a set that is different depending on the different devices. The net result would be that one person could administer a significantly greater amount of storage. An historical example exists with the mainframe operating system OS/390 where a consistent set of tools has been used for some time to administer storage at the system level. In that very complex environment, storage administrators can typically administer greater amounts than in an environment with multiple servers, operating systems and device types. Being able to administer more storage and to be more quickly trained on a specific, common set of management tools helps address the availability problem of people skilled in storage administration. The amount of additional storage that can be administered has not been quantified but different vendors are quoting amounts based on implementations of system-wide virtualization in a Storage Area Network that range from 4 to 40 times the amount of storage based on the types of solutions. These multipliers, when used against the actual costs of administration, provide the basis for tremendous business value to customers. Indeed, the vendors that have realized this have compelling sales arguments for their solutions.

Dealing with disk storage as a virtualized entity allows for new technology in the form of infrastructure or devices to be introduced without affecting the basic operation of the applications and systems as long as the virtualization method has not been changed. This greatly reduces the risk of deploying new technologies and

## **Virtualization of Disk Storage (WP-0007-1)**

provides the vehicle to realize the savings or advantages that the new technology should deliver. A disruptive technology where the basic operation must change would face significant business resistance to implementation. In that vein, traditional IT shops typically introduce infrastructure changes on a very infrequent basis. Depending on the type of environment, that change cycle could range from five to ten years because of the disruption, cost, and risks. With virtualization, a new approach could be taken to accommodate valuable new technologies.

Defining an abstraction layer for disk storage, which is the definition of virtualization in this case, presents a stratification of the complexity. The visibility of the complexity of the storage attached should be masked from the systems and from the day-to-day administration tasks.

### **Where Virtualization Should Be Done**

There are several choices for implementing virtualization and vendors have implemented a variety of solutions. Understanding the possible implementations and what is involved in doing and using the implementation should lead to a clear choice of where virtualization should be done. As virtualization has evolved, so have the expectations.

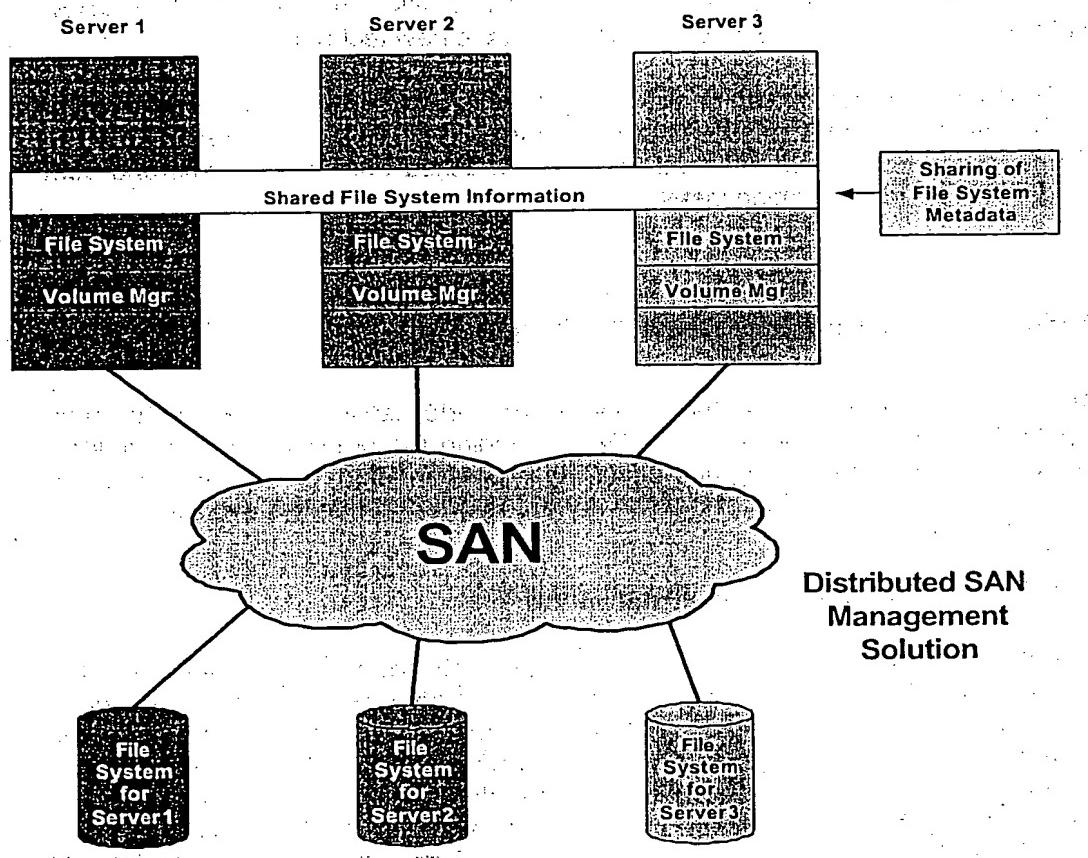
#### **Possible implementations of virtualization**

Individual devices can be virtualized within the disk system. This allows for self-contained virtualization where the logical device images presented to the interfaces (and therefore the systems attached) are mapped to physical devices in the storage system itself. The mapping can be done algorithmically or through a set of tables. Advanced functionality can be exploited in the storage system by special purpose software on all attached systems. The exploitation can be the dynamic creation and deletion of logical device images, point-in-time copies, and movement of data for balancing of performance / capacity. The exploitation software would be specific to that device and the administration would also be device specific.

A layer in the program stack could provide a virtualization of the physical devices attached to a server. The layer would be software to map the physical devices to an administrator specified set of logical devices. The mapping software would execute with each access to resolve location and be under total control of the administrator. This solution would result in a virtualization of storage for that particular server. Other servers could have the same layer in their program stacks. Unless specific software is installed on all attached servers and controlled by an administrator, the virtualization would not have a presentation of commonality across different systems. If the software to control virtualization were installed on all systems, it

## **Virtualization of Disk Storage (WP-0007-1)**

would have to share information (under control of an administrator) and maintain coherency and consistency across the potentially dissimilar systems to be functional. Figure 1 is a graphical representation of one type of a distributed storage solution.

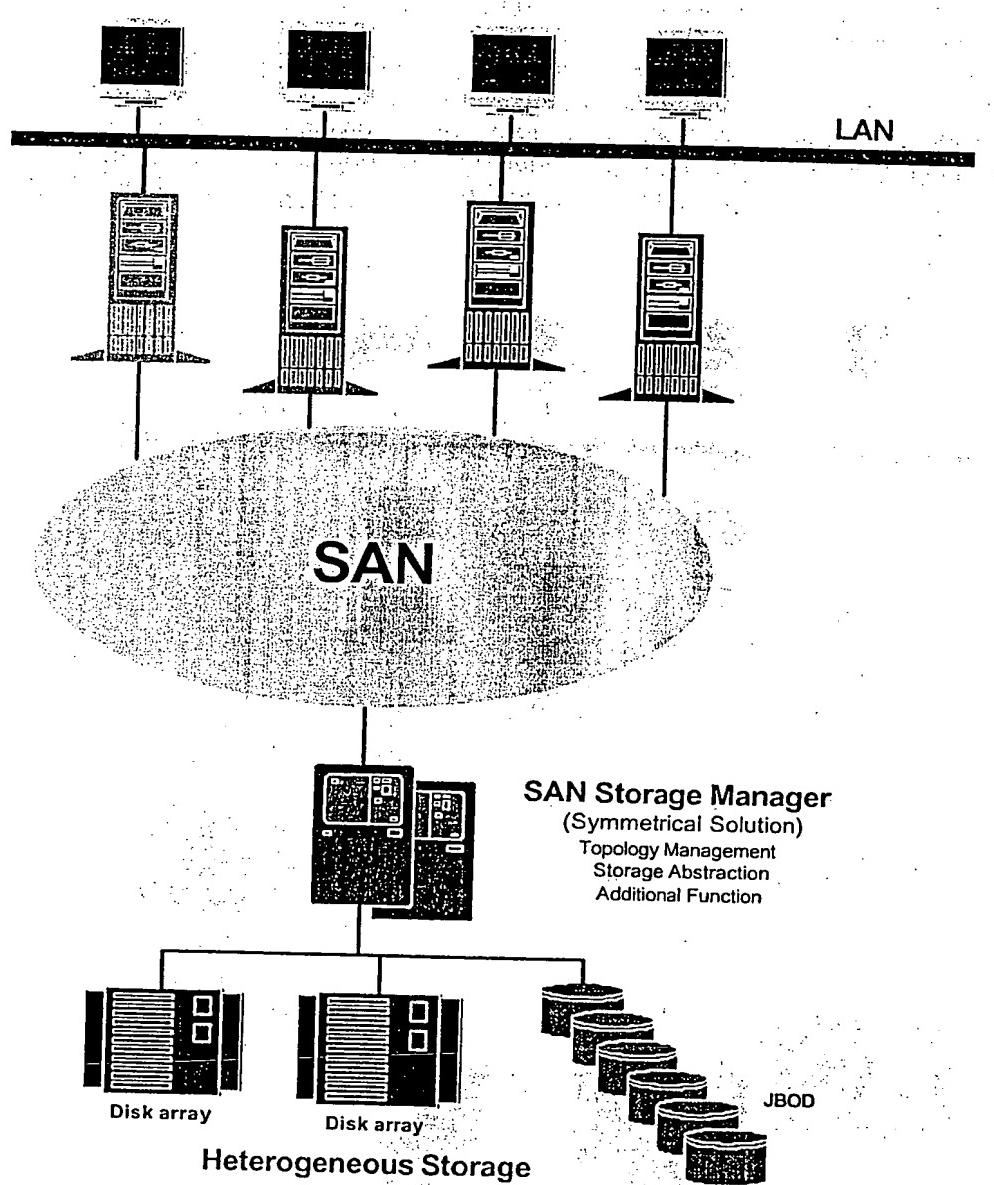


**Figure 1: A distributed SAN management solution**

Another approach is to use a device 'in line' for all accesses to storage that provides the abstraction of storage (as well as potentially other advanced functions). This would require that all systems route I/O access through this device (generically termed a SAN Storage Manager and called a symmetrical SAN solution) to have access to data. The SAN Storage Manager would handle the mapping of the logical representation of disk storage that it presents to attached systems to the physical devices attached to it. This eliminates the coherency and complexity issues by having a single entity control. With appropriate failover mechanisms, this provides the benefits of virtualization at the expense of another intelligent device in

## **Virtualization of Disk Storage (WP-0007-1)**

the data stream. Figure 2 is a graphical representation of a SAN Storage Manager solution.



**Figure 2: SAN Storage Manager**

## Virtualization of Disk Storage (WP-0007-1)

Yet another approach is a solution to provide the virtualization from a central control point but still allow I/O to be done directly to devices is another approach. In this case, cooperative software in the servers is used to obtain the information about the logical devices they will have access to and this central control (generically termed a "metadata server" and called an asymmetrical SAN solution) provides the information. The metadata server performs the mapping to provide the virtualization functions as well as the storage management tasks. In this case a common, centralized abstraction of storage is done with individual systems having an installed capability to obtain the logical device rather than accessing it directly from the physical device. Figure 3 is a graphical representation of a Metadata Server solution.

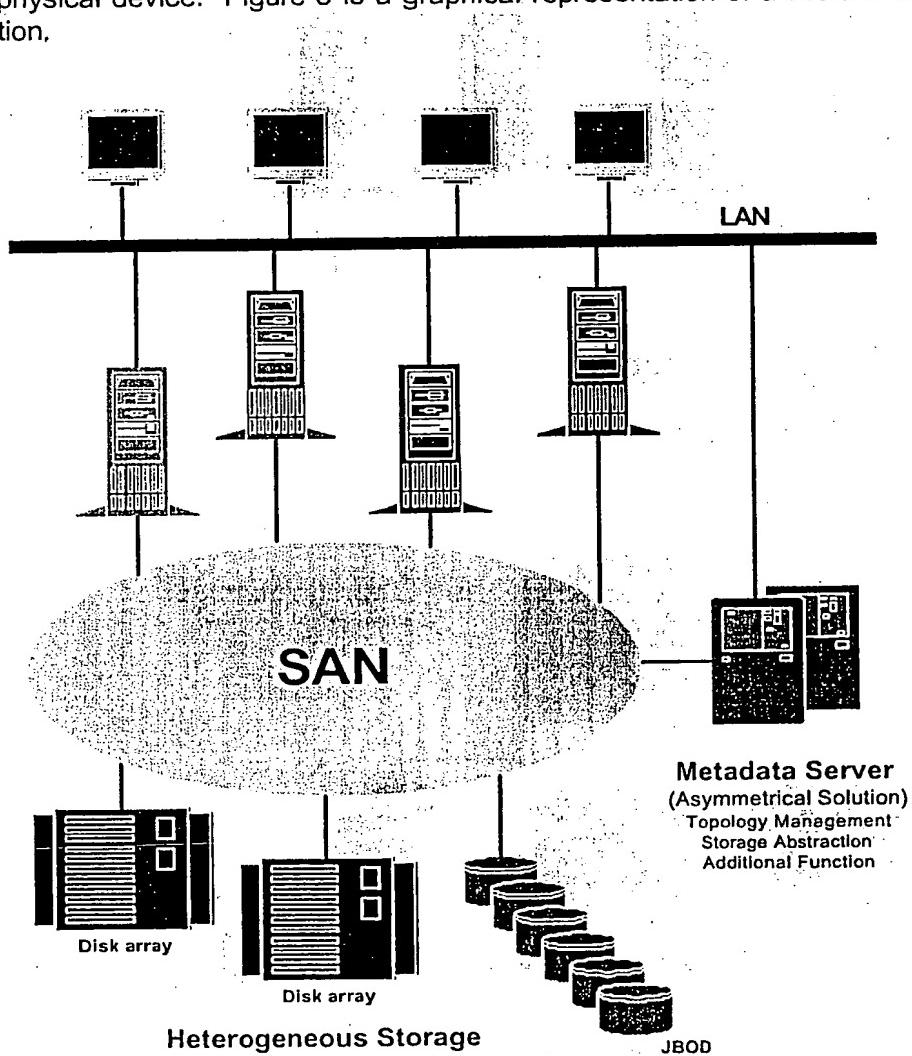


Figure 3: Metadata Server Solution

## **Virtualization of Disk Storage (WP-0007-1)**

### **Challenges to implement virtualization**

Implementing virtualization is a complex, software-intensive task. Assuring data integrity in the case where data has been moved or "redistributed" is critically important. Security of the access to data must also be maintained. Complexities also arise in the handling of error situations such as hardware failures.

An implementation that is done within a storage subsystem faces added complexities in that typically the code is executed on an embedded microprocessor with a special purpose operating system kernel of some type. The productivity for code development is much less and the testing characteristics more involved than development on a standard server and operating system. Add to that the extra hardware monitoring and control and the intricate error recovery, the development and successive enhancement of virtualization inside an individual storage system becomes a daunting challenge, with minimal exploitation capability.

Implementing virtualization as part of the program stack within a server represents a controlled environment requiring an intimate knowledge of the specific operating system and other components such as the file system. This activity is bounded within that server unless a distributed communication software element is added to cooperate among all servers attached to the target storage. Given that different systems (and varying levels of administration) may be involved, the potential for errors that may compromise integrity are much greater and require extensive testing and a very "controlled" administration.

Routing all disk storage I/O through a device such as a SAN Storage Manager or symmetric solution is effectively operating as a "high-function" disk controller. If the SAN Storage Manager is using a standard server platform and operating system, the development effort is less than a normal disk subsystem but is still faced with the complicated error recovery scenarios. If there is added capability to offset the additional overhead in the data path, this adds to the implementation challenge to achieve virtualization.

Using a central control point, the metadata server or asymmetric solution, to provide the virtualization while not being in the data path provides a well-bounded implementation challenge. Error recovery is limited here because server operating systems and the disk systems would handle errors the same as if there were no virtualization. The requirement for fault tolerant operation of the metadata server is still there but would probably be standard cluster operation. The one additional element is the indirection installed on each server to obtain the virtualized

## **Virtualization of Disk Storage (WP-0007-1)**

information about the logical devices. This would be isolated to each server but requires complete recovery, integrity, and security attention when implementing. The metadata server approach is also referred to as an asymmetric storage approach.

### **Complexity with multiple virtualization solutions**

The different types of implementation for virtualization have been discussed as if each is independently deployed. There may be attempts to utilize more than one solution in the same environment. Doing this has several ramifications that need to be considered. The overall complexity factor (the complexity of the different implementations) is not additive but is multiplied. The mapping of data to achieve virtualization becomes extremely complex when considering error recovery. Additionally it may become counter productive in the sense that the virtualization being performed by one solution may distribute data to provide an equal pattern of access among multiple devices. When cascaded with another solution that would be attempting the same technique it may have the exact opposite effect – artificial hot spots may be introduced thus having an adverse effect on overall performance. Utilizing multiple virtualization solutions may also have the effect of introducing an unacceptable amount of overhead due to the handling of multiple indirections.

### **The best approach to implementing virtualization**

It should be obvious that a virtualization solution that spans across systems in the enterprise provides the greater gains in benefits provided by virtualization. Individual solutions will not provide the same business value to customers that an enterprise-wide solution is capable of doing. The solution needs to be comprehensive in providing abstracted storage for all servers connected in the Storage Area Network. The business value seen in reduced administrative costs and the realization of the benefits of virtualization should be able to be demonstrated to reaffirm the choice for the virtualization solution. A comprehensive solution that virtualizes disk storage across the enterprise rather than at a specific subsystem has the potential for the greatest gains in reduction in administrative costs and enabling of enterprise-wide features.

### **Deployment of Virtualization**

As stated before, there are virtualization solutions available today and more being delivered. The question about deployment really should be one of when does virtualization of disk storage become as common as that of virtual memory? The virtual memory example is interesting to look at now. Can server operating systems even be purchased with only real memory implementations? Eventually, virtualization of disk storage will be that common.

## **Virtualization of Disk Storage (WP-0007-1)**

### **Timeframe**

Storage Area Networks are the major exploiters of virtualization. SANs are being deployed in increasing numbers and it is a logical next step to embrace a virtualization solution with a SAN to achieve the gains (value) beyond those of connectivity (the plumbing). By the end of 2001, half the major IT enterprises will have some type of SAN installed and a high percentage of those will have a virtualization solution.

The successes of these early implementations are expected to have in realizing the benefits of virtualization will probably accelerate the adoption by enterprise IT operations.

### **Choosing wisely**

Once implementing an enterprise-wide virtualization solution and achieving the benefits that brings, it would be very difficult to remove or even change the virtualization approach. Removal would require massive data movement and significant administrative tasks as well as giving up the advantages gained and introducing additional costs (and adding scarce people resources). Changing to a different virtualization approach is probably no less difficult in the administrative effort or massive data movement.

### **Competing implementations and standards**

There are different approaches being offered for virtualization and more are promised. All the approaches fit into the types outlined earlier. An "official" standard for virtualization from a standards body such as ANSI or ISO is unlikely and certainly not within a two year timeframe. Consequently, there will probably be more than one solution that will succeed in the market place. For the customer, the choices involve evaluating the business value of the approach, the ability of the vendor to deliver and support the solution, the vision and roadmap of the vendor's capabilities with regard to bringing additional value through functionality, and the risks the customer is willing to take in deploying a new strategy for storage management.

### **Futures**

Implementing virtualization of disk storage opens up many new possibilities. There are companies that will develop new applications that are based on disk storage virtualization to provide capabilities not achievable currently.

## **Virtualization of Disk Storage (WP-0007-1)**

### **Other software tie-in**

Existing software products will be adapted to exploit virtualized disk storage. Products providing backup, HSM, remote copy, etc. will be optimized to work with virtual devices. Previously intrusive operations will be made transparent. Companies that deliver these capabilities first will achieve a competitive advantage.

### **All storage management keyed to virtualization**

When a customer takes advantage of the virtualization of disk storage, savings in storage administrative costs will drive the usage of management tools that utilize and exploit virtualization. The gains will lead to storage management based entirely on virtualized storage. Overall benefit to the customer will dictate that only software capable of taking advantage of virtualized storage will be purchased for storage management.

### **Mainframe inclusion**

It is a natural evolution for the S/390 to participate in an enterprise-wide Storage Area Network, to provide a consistent storage management capability including virtualization. S/390 customers increasingly need to administer non-S/390 systems. The ability to have the same set of storage management functions enables the S/390 storage management expertise to be applied throughout the enterprise and leverage DFSMS skills. This would be a compelling sales story for customers.

## **Conclusions**

Virtualization of disk storage will provide such compelling benefits that it will inevitably be a common implementation for computer systems. The advantages for enterprise-wide virtualization will lead to that implementation becoming dominant over time. Vendors that articulate the value to customers, that can demonstrate the savings in administrative costs, and that can deliver will have a decided advantage. Customers will see productivity gains and cost containment as their storage systems grow to meet demand.

### **Indirect results of virtualization**

The full effect of virtualization is not completely clear but some projections can be made based on the implementations:

For disk storage, function will likely shift from the disk storage system to the SAN virtualization layer. With the abstraction of the physical disks, the functions/feature sets that today might differentiate one disk storage system

## **Virtualization of Disk Storage (WP-0007-1)**

from another will be exploited by software possibly running on a specific purpose appliances. Storage devices may have capabilities that the software can use to assist in the function, but the operating systems and applications will have the functionality regardless. The result will be that the customer will have flexibility in purchasing decisions and competitive pricing.

Storage management software will become a critical purchasing decision. The capability to utilize the virtualized environment in a manner that provides the greatest value will be the battleground for vendors. More analysis will be focused on these purchasing choices and more companies will be competing with solutions.

With the importance of storage being recognized more clearly today, strategic planning for storage strategies will become a critical need for companies. A new discipline will evolve that will first be filled by professional service organizations. This discipline will become a key component in information technology.

Virtualization of disk storage is the key enabler for advances in managing storage and accommodating growth to meet business demands. Only a complete enterprise-wide approach to virtualization of disk storage can provide the maximum benefits. Choosing the best approach and the vendor that has the best solutions will be a decision of critical importance for IT management.

**FEET TRANSMITTAL  
for FY 2005**

Effective 01/01/2004. Patent fees are subject to annual revision.

<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27	<b>TOTAL AMOUNT OF PAYMENT (\$)</b> <b>130.00</b>
--	---

<i>Complete if Known</i>	
Application Number	10/656,096
Filing Date	September 5, 2003
First Named Inventor	Idei, Hideomi
Examiner Name	Unassigned
Art Unit	2141
Attorney Docket No.	16869S-094000US

METHOD OF PAYMENT (check all that apply)				FEE CALCULATION (continued)																																																																																																																																																																				
<input type="checkbox"/> Check	<input type="checkbox"/> Credit Card	<input type="checkbox"/> Money Order	<input type="checkbox"/> Other	<input type="checkbox"/> None	3. ADDITIONAL FEES																																																																																																																																																																			
<input checked="" type="checkbox"/> Deposit Account:				<table border="1"> <thead> <tr> <th>Large Entity</th> <th>Small Entity</th> <th colspan="2">Fee Description</th> <th>Fee Paid</th> </tr> </thead> <tbody> <tr> <td>Fee Code</td> <td>Fee (\$)</td> <td>Fee Code</td> <td>Fee (\$)</td> <td></td> </tr> <tr> <td>1051</td> <td>130</td> <td>2051</td> <td>65</td> <td>Surcharge - late filing fee or oath</td> </tr> <tr> <td>1052</td> <td>50</td> <td>2052</td> <td>25</td> <td>Surcharge - late provisional filing fee or cover sheet.</td> </tr> <tr> <td>1053</td> <td>130</td> <td>1053</td> <td>130</td> <td>Non-English specification</td> </tr> <tr> <td>1812</td> <td>2,520</td> <td>1812</td> <td>2,520</td> <td>For filing a request for <i>ex parte</i> reexamination</td> </tr> <tr> <td>1804</td> <td>920*</td> <td>1804</td> <td>920*</td> <td>Requesting publication of SIR prior to Examiner action</td> </tr> <tr> <td>1805</td> <td>1,840*</td> <td>1805</td> <td>1,840*</td> <td>Requesting publication of SIR after Examiner action</td> </tr> <tr> <td>1251</td> <td>110</td> <td>2251</td> <td>55</td> <td>Extension for reply within first month</td> </tr> <tr> <td>1252</td> <td>430</td> <td>2252</td> <td>215</td> <td>Extension for reply within second month</td> </tr> <tr> <td>1253</td> <td>980</td> <td>2253</td> <td>490</td> <td>Extension for reply within third month</td> </tr> <tr> <td>1254</td> <td>1,530</td> <td>2254</td> <td>765</td> <td>Extension for reply within fourth month</td> </tr> <tr> <td>1255</td> <td>2,080</td> <td>2255</td> <td>1,040</td> <td>Extension for reply within fifth month</td> </tr> <tr> <td>1401</td> <td>340</td> <td>2401</td> <td>170</td> <td>Notice of Appeal</td> </tr> <tr> <td>1402</td> <td>340</td> <td>2402</td> <td>170</td> <td>Filing a brief in support of an appeal</td> </tr> <tr> <td>1403</td> <td>300</td> <td>2403</td> <td>150</td> <td>Request for oral hearing</td> </tr> <tr> <td>1451</td> <td>1,510</td> <td>1451</td> <td>1,510</td> <td>Petition to institute a public use proceeding</td> </tr> <tr> <td>1452</td> <td>110</td> <td>2452</td> <td>55</td> <td>Petition to revive – unavoidable</td> </tr> <tr> <td>1453</td> <td>1,330</td> <td>2453</td> <td>665</td> <td>Petition to revive – unintentional</td> </tr> <tr> <td>1501</td> <td>1,370</td> <td>2501</td> <td>685</td> <td>Utility issue fee (or reissue)</td> </tr> <tr> <td>1502</td> <td>490</td> <td>2502</td> <td>245</td> <td>Design issue fee</td> </tr> <tr> <td>1503</td> <td>660</td> <td>2503</td> <td>330</td> <td>Plant issue fee</td> </tr> <tr> <td>1460</td> <td>130</td> <td>1460</td> <td>130</td> <td>Petitions to the Commissioner</td> </tr> <tr> <td>1807</td> <td>50</td> <td>1807</td> <td>50</td> <td>Processing fee under 37 CFR 1.17(q)</td> </tr> <tr> <td>1806</td> <td>180</td> <td>1806</td> <td>180</td> <td>Submission of Information Disclosure Stmt</td> </tr> <tr> <td>8021</td> <td>40</td> <td>8021</td> <td>40</td> <td>Recording each patent assignment per property (times number of properties)</td> </tr> <tr> <td>1809</td> <td>790</td> <td>2809</td> <td>395</td> <td>Filing a submission after final rejection (37 CFR § 1.129(a))</td> </tr> <tr> <td>1810</td> <td>790</td> <td>2810</td> <td>395</td> <td>For each additional invention to be examined (37 CFR § 1.129(b))</td> </tr> <tr> <td>1801</td> <td>790</td> <td>2801</td> <td>395</td> <td>Request for Continued Examination (RCE)</td> </tr> <tr> <td>1802</td> <td>900</td> <td>1802</td> <td>900</td> <td>Request for expedited examination of a design application</td> </tr> <tr> <td colspan="4">Other fee (specify) _____</td> <td></td> </tr> <tr> <td colspan="4">*Reduced by Basic Filing Fee Paid</td> <td><b>SUBTOTAL (3)</b></td> <td><b>(\$130.00)</b></td> </tr> </tbody> </table>				Large Entity	Small Entity	Fee Description		Fee Paid	Fee Code	Fee (\$)	Fee Code	Fee (\$)		1051	130	2051	65	Surcharge - late filing fee or oath	1052	50	2052	25	Surcharge - late provisional filing fee or cover sheet.	1053	130	1053	130	Non-English specification	1812	2,520	1812	2,520	For filing a request for <i>ex parte</i> reexamination	1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action	1251	110	2251	55	Extension for reply within first month	1252	430	2252	215	Extension for reply within second month	1253	980	2253	490	Extension for reply within third month	1254	1,530	2254	765	Extension for reply within fourth month	1255	2,080	2255	1,040	Extension for reply within fifth month	1401	340	2401	170	Notice of Appeal	1402	340	2402	170	Filing a brief in support of an appeal	1403	300	2403	150	Request for oral hearing	1451	1,510	1451	1,510	Petition to institute a public use proceeding	1452	110	2452	55	Petition to revive – unavoidable	1453	1,330	2453	665	Petition to revive – unintentional	1501	1,370	2501	685	Utility issue fee (or reissue)	1502	490	2502	245	Design issue fee	1503	660	2503	330	Plant issue fee	1460	130	1460	130	Petitions to the Commissioner	1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	1806	180	1806	180	Submission of Information Disclosure Stmt	8021	40	8021	40	Recording each patent assignment per property (times number of properties)	1809	790	2809	395	Filing a submission after final rejection (37 CFR § 1.129(a))	1810	790	2810	395	For each additional invention to be examined (37 CFR § 1.129(b))	1801	790	2801	395	Request for Continued Examination (RCE)	1802	900	1802	900	Request for expedited examination of a design application	Other fee (specify) _____					*Reduced by Basic Filing Fee Paid				<b>SUBTOTAL (3)</b>	<b>(\$130.00)</b>
Large Entity	Small Entity	Fee Description		Fee Paid																																																																																																																																																																				
Fee Code	Fee (\$)	Fee Code	Fee (\$)																																																																																																																																																																					
1051	130	2051	65	Surcharge - late filing fee or oath																																																																																																																																																																				
1052	50	2052	25	Surcharge - late provisional filing fee or cover sheet.																																																																																																																																																																				
1053	130	1053	130	Non-English specification																																																																																																																																																																				
1812	2,520	1812	2,520	For filing a request for <i>ex parte</i> reexamination																																																																																																																																																																				
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action																																																																																																																																																																				
1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action																																																																																																																																																																				
1251	110	2251	55	Extension for reply within first month																																																																																																																																																																				
1252	430	2252	215	Extension for reply within second month																																																																																																																																																																				
1253	980	2253	490	Extension for reply within third month																																																																																																																																																																				
1254	1,530	2254	765	Extension for reply within fourth month																																																																																																																																																																				
1255	2,080	2255	1,040	Extension for reply within fifth month																																																																																																																																																																				
1401	340	2401	170	Notice of Appeal																																																																																																																																																																				
1402	340	2402	170	Filing a brief in support of an appeal																																																																																																																																																																				
1403	300	2403	150	Request for oral hearing																																																																																																																																																																				
1451	1,510	1451	1,510	Petition to institute a public use proceeding																																																																																																																																																																				
1452	110	2452	55	Petition to revive – unavoidable																																																																																																																																																																				
1453	1,330	2453	665	Petition to revive – unintentional																																																																																																																																																																				
1501	1,370	2501	685	Utility issue fee (or reissue)																																																																																																																																																																				
1502	490	2502	245	Design issue fee																																																																																																																																																																				
1503	660	2503	330	Plant issue fee																																																																																																																																																																				
1460	130	1460	130	Petitions to the Commissioner																																																																																																																																																																				
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)																																																																																																																																																																				
1806	180	1806	180	Submission of Information Disclosure Stmt																																																																																																																																																																				
8021	40	8021	40	Recording each patent assignment per property (times number of properties)																																																																																																																																																																				
1809	790	2809	395	Filing a submission after final rejection (37 CFR § 1.129(a))																																																																																																																																																																				
1810	790	2810	395	For each additional invention to be examined (37 CFR § 1.129(b))																																																																																																																																																																				
1801	790	2801	395	Request for Continued Examination (RCE)																																																																																																																																																																				
1802	900	1802	900	Request for expedited examination of a design application																																																																																																																																																																				
Other fee (specify) _____																																																																																																																																																																								
*Reduced by Basic Filing Fee Paid				<b>SUBTOTAL (3)</b>	<b>(\$130.00)</b>																																																																																																																																																																			
<b>SUBTOTAL (1)</b> <b>(\$0.00)</b>																																																																																																																																																																								
<b>SUBTOTAL (2)</b> <b>(\$0.00)</b>																																																																																																																																																																								
**or number previously paid, if greater; For Reissues, see above																																																																																																																																																																								

SUBMITTED BY					
Name (Print/Type)	Chun-Pok Leung	Registration No. (Attorney/Agent)	41,405	Telephone	650-326-2400
Signature	<i>Chun-Pok Leung</i>			Date	December 17, 2004

**WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.**